

# On the Number of Fixed-Length Semiorders

Yangzhou Hu  
Massachusetts Institute of Technology

July 17, 2012

## Abstract

A *semiorder* is a special type of partially ordered set. We find a bijection between  $n$ -element semiorders with length  $H$  and  $(n + 1)$ -node ordered trees of height  $H + 2$ . This bijection preserves not only the number of elements, but also much additional structure. Based on this correspondence, we calculate the generating functions and explicit formulas for the numbers of labeled and unlabeled  $n$ -element semiorders of length  $H$ . We also prove several concise recurrence relations and provide combinatorial proofs for special cases of the explicit formulas.

**Keywords:** semiorder, ordered tree, Dyck path,

## 1 Introduction and Main Theorem

A **semiorder** is a partially ordered set without the following two structures:

- **(2 + 2)**-structure: four distinct elements  $x, y, z, w$ , such that  $x > y, z > w$ , and other pairs are incomparable.
- **(3 + 1)**-structure: four distinct elements  $x, y, z, w$ , such that  $x > y > z$ , and other pairs are incomparable.

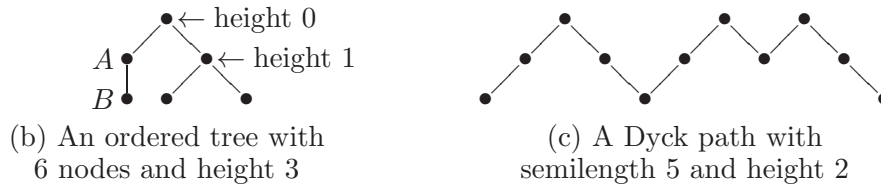


In other words, semiorders are **(2 + 2)**-structure-free and **(3 + 1)**-structure-free posets, or just **(2 + 2)**-free and **(3 + 1)**-free. Every semiorder can also be regarded as a partial ordering  $P$  of a subset of  $\mathbb{R}$  defined by  $x < y$  in  $P$  if  $x < y - 1$  in  $\mathbb{R}$ . The length  $H$  of a semiorder is the length of a longest chain. Every semiorder  $R$  with  $n$  elements, up to isomorphism, can be uniquely represented as an integer vector  $\rho(R) = (r_1, r_2, \dots, r_n)$ , where  $r_i$  represents the number of elements smaller than the  $i^{\text{th}}$  element, and  $r_1 \geq r_2 \geq \dots \geq r_n \geq 0$ ,  $r_i \leq n - i$ , for all  $1 \leq i \leq n$ . For instance, the above graph (a) presents a semiorder  $R$  with 5 elements, length 1, and vector  $\rho(R) = (3, 2, 0, 0, 0)$ . For further basic information on semiorders, see [3].

There is much interest in enumerating the number of posets with certain properties. For example, Kitaev and Rémel enumerated the number of **(2 + 2)**-free posets [4]. It is a classical result of Wine and Freund [6] that the number of nonisomorphic  $n$ -element semiorders is the Catalan number  $C_n = \frac{1}{n+1} \binom{2n}{n}$  and Chandon, Lemaire, and Pouget [2] showed (in an equivalent form) that

if  $f(n)$  is the number of  $n$ -element labelled semiorders (or semiorders on an  $n$ -element set), then  $\sum_{n \geq 0} f(n) \frac{x^n}{n!} = \sum_{n \geq 0} C_n (1 - e^{-x})^n$ . For a general principle implying this result, see Lemma 3.4. In this paper, we deal with semiorders of length at most  $H$ . That is, we enumerate the number of posets which are  $(\mathbf{2} + \mathbf{2})$ -free,  $(\mathbf{3} + \mathbf{1})$ -free, and of length at most  $H$ . We conduct the enumeration by establishing a bijection between semiorders and ordered trees of a fixed height.

An **ordered tree** is a rooted tree which has been embedded in the plane so that the relative order of subtrees at each node is part of its structure. The height  $H$  of an ordered tree is the number of nodes in a chain of maximum length. The following graph (b) presents an ordered tree with 6 nodes and height 3. Notice that in order to be consistent with much of the literature, we define the length of a semiorder based on the number of edges of its longest chain, while defining the height of an ordered tree based on the number of nodes of its longest chain.



A **Dyck path** with semilength  $n$  is a lattice path in the Euclidean plane from  $(0, 0)$  to  $(2n, 0)$  whose steps are either  $(1, 1)$  or  $(1, -1)$  and the path never goes below the  $x$ -axis. The height  $H$  of a Dyck path is the maximal  $y$ -coordinate among all points on the path. The above graph (c) shows a Dyck path with semilength 5 and height 2.

It is well-known that there is a one-to-one correspondence between (i) ordered trees with  $n + 1$  nodes and height  $H + 1$  and (ii) Dyck paths with semilength  $n$  and height  $H$ . This paper establishes a bijection between  $n$ -element semiorders of length  $H$  and  $(n + 1)$ -node ordered trees of height  $H + 2$ . Thus these semiorders simultaneously correspond to Dyck paths of semilength  $n$  and height  $H + 1$ .

**Theorem 1.1** (Main Theorem). *For  $n \geq 1$  and  $H \geq 0$ , the number of nonisomorphic  $n$ -element unlabeled semiorders of length  $H$  is equal to the number of  $(n + 1)$ -node ordered trees of height  $H + 2$ , which is also equal to the number of Dyck paths with semilength  $n$  and height  $H + 1$ .*

Section 2 gives a recurrence proof and a bijective proof for Theorem 1.1. Section 3 calculates the generating functions and explicit formulas for the number of unlabeled as well as labeled semiorders with fixed lengths. Section 4 proves some concise recurrence relations, and Section 5 shows explicit formulas for the number of semiorders with certain lengths  $H$ , and provides simple bijective proofs for these formulas.

## 2 Proof of Main Theorem 1.1

Before proving Theorem 1.1, we first define some terminology that is used later in the proof.

**Definition 2.1.** A node  $A$  of an ordered tree has height  $i$  if the distance from node  $A$  to the root is  $i$ . In particular, the root has height 0.

In this paper we regard the root as the uppermost node, and all other nodes are below the root. We say node  $B$  is attached to node  $A$  if node  $B$  has height  $i + 1$  and node  $A$  has height  $i$ , and these two nodes are adjacent. Refer to graph (b) as an example.

**Definition 2.2.** An element  $a$  of a semiorder is on the  $i^{\text{th}}$  level ( $i \geq 1$ ) if  $i$  is the largest integer for which there exist  $i - 1$  elements  $a_1, a_2, \dots, a_{i-1}$  satisfying  $a_1 > a_2 > \dots > a_{i-1} > a$ . Refer to graph (a) as an example.

**Proposition 2.3.** *For a length  $H$  semiorder, and for  $1 \leq i \leq H$ , there is at least one element on the  $i^{\text{th}}$  level that is larger than all elements on the  $(i+1)^{\text{th}}$  level.*

*Proof.* Suppose that there does not exist an element on the  $i^{\text{th}}$  level that is larger than all elements on the  $(i+1)^{\text{th}}$  level. Since every element on the  $(i+1)^{\text{th}}$  level must be smaller than at least one element on the  $i^{\text{th}}$  level, there must exist two elements  $a$  and  $c$  on the  $(i+1)^{\text{th}}$  level which are smaller than two distinct elements  $b$  and  $d$  on the  $i^{\text{th}}$  level, respectively, and  $b$  is not larger than  $c$ , while  $d$  is not larger than  $a$ . Then  $\{b > a, d > c\}$  forms a  $(\mathbf{2} + \mathbf{2})$ -structure, a contradiction. Therefore, at least one element on the  $i^{\text{th}}$  level is larger than all elements on the  $(i+1)^{\text{th}}$  level.  $\square$

**Proposition 2.4.** *For a length  $H$  semiorder, and for  $1 \leq i < j \leq H+1$ ,  $j-i \geq 2$ , every element on the  $i^{\text{th}}$  level is larger than all elements on the  $j^{\text{th}}$  level.*

*Proof.* Assume to the contrary that there exist an element  $b$  on the  $i^{\text{th}}$  level and an element  $a$  on the  $j^{\text{th}}$  level such that  $b$  is not larger than  $a$ . By Definition 2.2, there exist  $j-1$  elements  $a_1, a_2, \dots, a_{j-1}$ , such that  $a_1 > a_2 > \dots > a_{j-1} > a$ , and then  $a_{j-2}$  and  $a_{j-1}$  should be on the  $(j-2)^{\text{th}}$  and  $(j-1)^{\text{th}}$  level, respectively. Since element  $b$  is on the  $i^{\text{th}}$  level, and  $j-i \geq 2$ , element  $b$  cannot be smaller than any of  $a_{j-2}, a_{j-1}$ , or  $a$ . In addition, since  $b$  is not larger than  $a$  and  $a_{j-2} > a_{j-1} > a$ ,  $b$  is not larger than any of  $a_{j-2}, a_{j-1}$  or  $a$ . Hence,  $\{a_{j-2} > a_{j-1} > a, b\}$  forms a  $(\mathbf{3} + \mathbf{1})$ -structure, a contradiction. Therefore every element on the  $i^{\text{th}}$  level should be larger than all elements on the  $j^{\text{th}}$  level, for  $j-i \geq 2$ .  $\square$

We are now ready to prove the Main Theorem 1.1. We give two proofs here: one considers the recurrence formulas of the two numbers in the theorem, and the other directly establishes a bijection between semiorders and ordered trees.

## 2.1 Recurrence proof

Let  $t(n, h, k)$  denote the number of  $(n+1)$ -node ordered trees of height  $h+2$ , and exactly  $k$  nodes have height  $h+1$ ,  $1 \leq k \leq n$ . Let  $f(n, h, k)$  be the number of  $n$ -element semiorders of length  $h$ , and exactly  $k$  elements are on the last level. We show that  $t(n, h, k)$  and  $f(n, h, k)$  have the same initial value and recurrence formula in the following lemmas, and thus they are equal.

**Lemma 2.5.** *For  $h \geq 1$ , we have*

$$t(n, h, k) = \sum_{m=1}^{n-k} \binom{m+k-1}{m-1} \cdot t(n-k, h-1, m). \quad (2.1)$$

*Proof.* Say we have an  $(n-k+1)$ -node ordered tree of height  $h+1$ , and assume that exactly  $m$  nodes have height  $h$ ,  $1 \leq m \leq n-k$ . Consider adding  $k$  nodes to the tree to get a new tree with  $n+1$  nodes and height  $h+2$ , and the newly added nodes are exactly the set of nodes of height  $h+2$ . Thus we need to attach the  $k$  new nodes to the  $m$  nodes of height  $h+1$ , and every new node is uniquely attached to one node. Let the  $m$  nodes be  $A_1, A_2, \dots, A_m$ , and the number of new nodes attached to  $A_i$  be  $r_i$ ,  $1 \leq i \leq m$ . Then we have  $r_1 + r_2 + \dots + r_m = k$ ,  $r_i \geq 0$ ,  $1 \leq i \leq m$ .

The number of integer solutions to the above equation is  $\binom{m+k-1}{m-1}$ . Therefore, we have  $\binom{m+k-1}{m-1}$  ways to add the  $k$  nodes. Summing up all possible  $m$ 's, we obtain

$$t(n, h, k) = \sum_{m=1}^{n-k} \binom{m+k-1}{m-1} \cdot t(n-k, h-1, m).$$

$\square$

**Lemma 2.6.** *For  $h \geq 1$ , we have*

$$f(n, h, k) = \sum_{m=1}^{n-k} \binom{m+k-1}{m-1} \cdot f(n-k, h-1, m). \quad (2.2)$$

*Proof.* We say that an element of a semiorder is *good* if the element is on the last level of the semiorder. Say we have an  $(n-k)$ -element semiorder  $S$  of length  $h-1$  and  $m$  good elements,  $1 \leq m \leq n-k$ . Consider adding  $k$  elements to  $S$  to get a new semiorder  $S'$  with  $n$  elements and length  $h$ , and the newly added elements are exactly the set of good elements of  $S'$ . Call the original  $m$  good elements  $a_1, a_2, \dots, a_m$ , and the  $k$  new elements  $b_1, b_2, \dots, b_k$ . Then in the semiorder  $S'$ , we have that  $a_1, a_2, \dots, a_m$  are the only elements on the  $h^{th}$  level, and  $b_1, b_2, \dots, b_k$  are the only elements on the  $(h+1)^{th}$  level.

If we remove all elements on the first  $h-1$  levels of  $S'$ , then we get a length one semiorder  $P$  with  $m+k$  elements, and there are exactly  $m$  elements on the upper level and  $k$  elements on the lower level. On the other hand, given a semiorder  $S$  and a semiorder  $P$  as above, we can uniquely determine the semiorder  $S'$ , because based on Proposition 2.4, the  $k$  elements on the  $(h+1)^{th}$  level of  $S'$  must be smaller than all elements on the  $i^{th}$  level in  $S'$ , for  $1 \leq i \leq h-1$ . Therefore, the semiorder  $P$  uniquely determines the way to add the  $k$  new elements.

Let  $P$  with  $\rho(P) = (p_1, p_2, \dots, p_{m+k})$  represent one such semiorder. Then we have

$$\begin{cases} k = p_1 \geq p_2 \geq \dots \geq p_m \geq 0 \\ p_{m+1} = p_{m+2} = \dots = p_{m+k} = 0. \end{cases} \quad (2.3)$$

Notice that  $\{p_2, p_3, \dots, p_m\}$  is an  $(m-1)$ -element multiset with elements from  $\{0, 1, \dots, k\}$ , and thus we have  $\binom{k+1+m-1-1}{m-1} = \binom{m+k-1}{m-1}$  such multisets. Therefore, there are  $\binom{m+k-1}{m-1}$  possible semiorder  $P$ 's. Summing up all possible  $m$ 's, we have

$$f(n, h, k) = \sum_{m=1}^{n-k} \binom{m+k-1}{m-1} \cdot f(n-k, h-1, m).$$

□

*Proof of the Main Theorem 1.1.* For  $h = 0$ , the  $(n+1)$ -element ordered tree of height  $h+2 = 2$  can only be the tree with  $n$  nodes adjacent to the root; meanwhile, the  $n$ -element semiorder of length 0 can only be the one with  $n$  elements and any two of the elements are incomparable. As a result, we have

$$t(n, 0, k) = f(n, 0, k) = \begin{cases} 1 & \text{if } n = k \\ 0 & \text{if } n \neq k. \end{cases}$$

For  $h \geq 1$ , by Lemma 2.5 and 2.6,  $t(n, h, k)$  and  $f(n, h, k)$  have the same recurrence formulas. Therefore  $t(n, h, k) = f(n, h, k)$  for every  $n \geq 1$ ,  $h \geq 0$ , and  $1 \leq k \leq n$ . Summing on  $k$  completes the proof of Theorem 1.1. □

## 2.2 Bijective proof

Recall that an element of a semiorder is good if it is on the last level of the semiorder. Based on the idea in the recurrence proof, we can construct a bijection between  $n$ -element semiorders of length

$H$  with  $k$  good elements and  $(n + 1)$ -element ordered trees of height  $H + 2$  with  $k$  nodes of height  $H + 1$ .

An ordered tree with  $n + 1$  nodes and height  $H + 2$  has  $H + 1$  levels (not including the  $0^{th}$  level). Assume there are  $x_i$  nodes on the  $i^{th}$  level,  $0 \leq i \leq H + 1$ . Since the root is the only node on the  $0^{th}$  level, we have

$$\sum_{i=1}^{H+1} x_i = n. \quad (2.4)$$

Let  $s_j^i$  denote the number of nodes on the  $i^{th}$  level that are adjacent to the  $j^{th}$  node on the  $(i - 1)^{th}$  level,  $1 \leq j \leq x_{i-1}$ ,  $1 \leq i \leq H + 1$ . Since every node on the  $i^{th}$  level should be adjacent to exactly one node on the  $(i - 1)^{th}$  level, we must have

$$\sum_{j=1}^{x_{i-1}} s_j^i = x_i. \quad (2.5)$$

Let  $u_j^i = \sum_{k=j}^{x_{i-1}} s_k^i$ ,  $1 \leq j \leq x_{i-1}$ ,  $1 \leq i \leq H + 1$ . Then  $u_1^i \geq u_2^i \geq \dots \geq u_{x_{i-1}}^i$ . Let  $y_i = \sum_{k=1}^i x_k$ ,  $1 \leq i \leq H + 1$ , and  $y_0 = 0$ . We now define  $R^i$  by induction, and let the number of entries in  $R^i$  be  $y_i$ .

Set  $R^1 = (0, 0, \dots, 0)$ , in which there are  $x_1 = y_1$  zeros. Assume  $R^i = (r_1^i, r_2^i, \dots, r_{y_i}^i)$ ,  $1 \leq i \leq H$ , and let

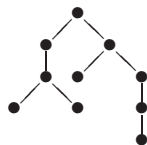
$$R^{i+1} = (r_1^i + x_{i+1}, r_2^i + x_{i+1}, \dots, r_{y_{i-1} + x_{i+1}}^i + x_{i+1}, r_{y_{i-1} + x_{i+1} + u_1^{i+1}}^i, r_{y_{i-1} + x_{i+1} + u_2^{i+1}}^i, \dots, r_{y_{i-1} + x_{i+1} + u_{x_i}^{i+1}}^i, 0, 0, \dots, 0)$$

in which there are  $x_{i+1}$  zeros, and thus  $y_{i-1} + x_i + x_{i+1} = y_{i+1}$  entries.

**Theorem 2.7.** *The vector  $R^{H+1}$  represents an  $n$ -element semiorder of length  $H$  with  $x_{H+1}$  good elements. This gives a bijective map from  $(n + 1)$ -node ordered trees of height  $H + 2$  to  $n$ -element semiorders of length  $H$ .*

Since this map is naturally derived from the recurrence proof, we do not give a rigorous proof on why the map is valid and why it is a bijection. The main idea here is to map an ordered tree of height  $H + 2$  to a semiorder with  $H + 1$  levels, where the number of elements on the  $i^{th}$  level in the semiorder is equal to the number of nodes of height  $i$  in the tree,  $1 \leq i \leq H + 1$ . We set a bijection between (a) the connections between nodes of heights  $i$  and  $i + 1$  in the tree, and (b) the set of ordered pairs between elements on level  $i$  and  $i + 1$  in the semiorder,  $1 \leq i \leq H$ . This bijection preserves not only the number of elements but also much additional structure. It presents an effective way to connect semiorders and ordered trees, as well as Dyck paths. In order to illustrate the bijection more clearly, we show by an example how the map works.

**Example 2.8.** Assume we have the following ordered tree:



Here the number of nodes is 10 and height is 5. It should correspond to a semiorder with 9 elements, length 3, and the number of elements on each level given by  $(2, 3, 3, 1)$ . Write  $s^i =$

$(s_1^i, s_2^i, \dots, s_{x_{i-1}}^i)$  and  $u^i = (u_1^i, u_2^i, \dots, u_{x_{i-1}}^i)$ . We have  $(x_1, x_2, x_3, x_4) = (2, 3, 3, 1)$ . Then the vectors representing the connections between two adjacent levels in the ordered tree are

$$s^1 = (2), \quad s^2 = (1, 2), \quad s^3 = (2, 0, 1), \quad s^4 = (0, 0, 1).$$

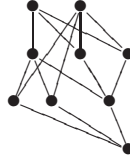
We transform these vectors into vectors that can represent the set of ordered pairs between two adjacent levels of the semiorder. These vectors are

$$u^1 = (2), \quad u^2 = (3, 2), \quad u^3 = (3, 1, 1), \quad u^4 = (1, 1, 1).$$

Now let us construct  $R^i$ ,  $1 \leq i \leq 4$ :

$$R^1 = (0, 0), \quad R^2 = (3, 2, 0, 0, 0), \quad R^3 = (6, 5, 3, 1, 1, 0, 0, 0), \quad R^4 = (7, 6, 4, 2, 2, 1, 1, 1, 0).$$

In fact,  $R^i$  depicts the semiorder with only the first  $i$  levels,  $1 \leq i \leq 4$ , and  $R^4$  is the final semiorder we desired. Its Hasse diagram is as follows:



The inverse map can be done by reversing the steps.

### 3 Generating Functions and Explicit Formulas

#### 3.1 On unlabeled semiorders

Let  $f_h^n$  denote the number of nonisomorphic unlabeled semiorders with  $n$  elements and length  $h$ , and  $f_{\leq h}^n$  denote the number of nonisomorphic unlabeled semiorders with  $n$  elements and length at most  $h$ , so  $f_h^n = f_{\leq h}^n - f_{\leq (h-1)}^n$ . Let  $F_h(x) = \sum_{n=0}^{\infty} f_h^n x^n$ , and  $F_{\leq h}(x) = \sum_{n=0}^{\infty} f_{\leq h}^n x^n$ .

De Bruijn, Knuth, and Rice [1] calculated the generating function for the number of fixed-height ordered trees in 1972. Based on this generating function and Theorem 1.1, we have the following corollary.

**Corollary 3.1.** For  $h \geq 0$ ,

$$\bullet F_h(x) = \sum_{n=0}^{\infty} f_h^n x^n = \frac{x^{h+1}}{p_{h+1}(x)p_h(x)} \quad (3.1)$$

$$\bullet F_{\leq h}(x) = \sum_{n=0}^{\infty} f_{\leq h}^n x^n = \frac{p_h(x)}{p_{h+1}(x)} \quad (3.2)$$

where

$$p_0(x) = 1, \quad p_1(x) = 1 - x, \quad p_{h+1}(x) = p_h(x) - x \cdot p_{h-1}(x).$$

De Bruijn, Knuth, and Rice [1] also found the explicit formulas for the number of fixed-height ordered trees. Based on their results and Theorem 1.1, we have the following corollary:

**Corollary 3.2.**  $f_{\leq h}^1 = f_{\leq h}^0 = 1$ . For  $n \geq 2$ ,  $h \geq 0$ , we have

$$f_{\leq h}^n = (h+3)^{-1} \sum_{1 \leq j \leq \frac{h+2}{2}} 4^{n+1} \sin^2 \left( \frac{j\pi}{h+3} \right) \cos^{2n} \left( \frac{j\pi}{h+3} \right). \quad (3.3)$$

### 3.2 On labeled semiorders

Let  $g_h^n$  denote the number of nonisomorphic labeled semiorders with  $n$  elements and length  $h$ , and  $g_{\leq h}^n$  denote the number of nonisomorphic labeled semiorders with  $n$  elements and length at most  $h$ . Thus  $g_h^n = g_{\leq h}^n - g_{\leq (h-1)}^n$ . Let  $G_h = \sum_{n=0}^{\infty} g_h^n \frac{x^n}{n!}$  and  $G_{\leq h} = \sum_{n=0}^{\infty} g_{\leq h}^n \frac{x^n}{n!}$ .

We obtain the exponential generating function  $G_h$  from the ordinary generating function  $F_h$  by the following lemma, which is due to Y. Zhang [7]. We first define equivalence of elements and then state the lemma.

**Definition 3.3.** Two elements  $p$  and  $p'$  of a poset  $P$  are *equivalent* if

$$p' < q \Leftrightarrow p < q, \text{ for all } q \in P$$

and

$$p' > q \Leftrightarrow p > q, \text{ for all } q \in P.$$

**Lemma 3.4.** Define the following two operations on an unlabeled poset  $P$ .

- The expansion of  $P$  at  $p \in P$  is obtained from  $P$  by adjoining a new element  $p'$  such that  $p$  and  $p'$  are equivalent.
- The contraction  $c(P)$  of  $P$  is a poset  $c(P)$  obtained from  $P$  by replacing every equivalence class of elements with a single element. Call a poset  $P$  a *seed* if  $P = c(P)$ . Call a seed  $P$  *rigid* if  $P$  has no nontrivial automorphisms.

Let  $C$  be a family of unlabeled posets such that  $C$  is closed under expansion and contraction, and all seeds in  $C$  are rigid. Let  $F(x) = \sum_{P \in C} x^{\#P}$  and  $G(x) = \sum_{P \in C} D_P \frac{x^{\#P}}{(\#P)!}$ , where  $\#P$  is the number of elements in poset  $P$  and  $D_P$  is the number of ways to label the elements of  $P$ . Then  $G(x) = F(1 - e^{-x})$ .

The class of semiorders of length  $h$  is closed under expansion and contraction. Zhang has observed that all  $(\mathbf{2}+\mathbf{2})$ -free seeds are rigid. Since semiorders are  $(\mathbf{2}+\mathbf{2})$ -free, all seeds of semiorders are rigid. As a result, Lemma 3.4 implies the following corollary.

**Corollary 3.5.** For  $h \geq 0$ ,

$$G_h(x) = F_h(1 - e^{-x}) = \frac{(1 - e^{-x})^{h+1}}{p_{h+1}(1 - e^{-x})p_h(1 - e^{-x})} \quad (3.4)$$

and

$$G_{\leq h}(x) = F_{\leq h}(1 - e^{-x}) = F_{\leq h}(1 - e^{-x}) = \frac{p_h(1 - e^{-x})}{p_{h+1}(1 - e^{-x})}. \quad (3.5)$$

## 4 Recurrence Relations

The generating functions and explicit formulas for the number of semiorders of fixed length are complicated, but we find some concise recurrence relations underneath. In this section, we provide simple combinatorial proofs of these relations. In this way, we can better understand the relations between fixed-length semiorders with different numbers of elements.

#### 4.1 Recurrence formula 1

**Theorem 4.1.** For  $n \geq 2$  and  $h \geq 1$ ,

$$f_{\leq h}^n = \sum_{t=0}^{n-1} f_{\leq h}^t f_{\leq h-1}^{n-1-t}, \quad (4.1)$$

where  $f_{\leq h}^0 = 1$ .

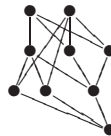
*Proof.* Say  $S$  is an  $n$ -element semiorder of length at most  $h$ . Say the integer vector corresponding to semiorder  $S$ , as discussed in Section 1, is  $(r_1, r_2, \dots, r_n)$ . For elements  $a$  and  $b$  on the same level of  $S$ , we say that element  $b$  is *to the right of* element  $a$  if  $b$  corresponds to  $r_j$ , while  $a$  corresponds to  $r_i$ ,  $i < j$ . Such an ordering is unique up to isomorphism.

Let  $a_1$  be the *rightmost* element on the first level of  $S$ , and let  $T_1 = \{a_1\}$ . Once  $T_i$  is defined, let  $T_{i+1}$  be the set of elements on the  $(i+1)^{th}$  level, each of whose elements is smaller than at least one element in  $T_i$ ,  $1 \leq i \leq h$ . For a given semiorder  $S$ , the set  $T_i$  is uniquely determined,  $1 \leq i \leq h$ . Notice that it is possible that  $T_i = \emptyset$ , for some  $i$ ,  $1 \leq i \leq h$ , and if  $T_i = \emptyset$ , then we must have  $T_j = \emptyset$  for all  $i \leq j \leq h+1$ . Let  $A_2 = T_1 \cup T_2 \cup \dots \cup T_{h+1}$ , and  $A_1 = A - A_2$ , where  $A$  is the set of all elements of  $S$ . Since  $a_1 \in A_2$ , we must have  $1 \leq |A_2| \leq n$ ,  $0 \leq |A_1| \leq n-1$ , and  $|A_1| + |A_2| = n$ .

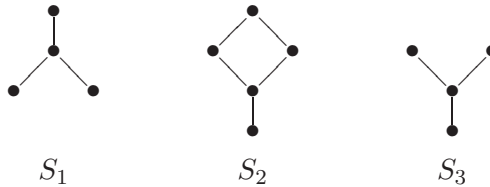
Let us separate  $S$  into two semiorders  $S_1$  and  $S_2$ . Let  $S_1$  be the induced semiorder with element set  $A_1$ . Similarly, let  $S_2$  be the induced semiorder with element set  $A_2$ . Let  $S_3$  be the semiorder obtained from  $S_2$  by removing element  $a_1$ . Then for a given semiorder  $S$ , we have that  $S_1$ ,  $S_2$ ,  $S_3$  are uniquely defined. Since  $S$  is a semiorder of length at most  $h$ , semiorders  $S_1$  and  $S_2$  have length at most  $h$ , and thus  $S_3$  has length at most  $h-1$ .

Assume  $|A_1| = t$ , so  $S_3$  is a semiorder with  $n-1-t$  elements. As a result, for a given  $n$ -element semiorder  $S$  of length at most  $h$ , we can uniquely obtain a pair of semiorders  $S_1$  and  $S_3$ , of length at most  $h$  and at most  $h-1$ , and with  $t$  and  $n-1-t$  elements, respectively,  $0 \leq t \leq n-1$ .

For example, if  $S$  is as follows,



the corresponding  $S_1$ ,  $S_2$ ,  $S_3$  are:



On the other hand, given a pair of semiorders  $S_1$  and  $S_3$ , of length at most  $h$  and at most  $h-1$ , and with  $t$  and  $n-1-t$  elements, respectively,  $0 \leq t \leq n-1$ , we can first add an element  $a_1$  to  $S_3$ , and let  $a_1$  be larger than all other elements in  $S_3$ . Let us call the new semiorder  $S_2$ . Then  $S_2$  has  $n-t$  elements and length at most  $h$ .

Let us construct a new semiorder  $S$  by combining  $S_1$  and  $S_2$  as follows. The elements on the  $i^{th}$  level of  $S$  are the elements on the  $i^{th}$  level of  $S_1$  and  $S_2$ ,  $1 \leq i \leq h+1$ , and the order relations in  $S_1$  and  $S_2$  are preserved. In addition, let every element on the  $(i-1)^{th}$  level of  $S_1$  be larger than all elements on the  $i^{th}$  level of  $S_2$ .



This construction uniquely gives an  $n$ -element semiorder  $S$  of length at most  $h$ , and if we separate  $S$  into two semiorders by the method discussed above, we get back  $S_1$  and  $S_3$ .

As a result, there is a one-to-one map between  $n$ -element semiorders  $S$  of length at most  $h$ , and pairs of semiorders  $S_1$  and  $S_3$ , of length  $\leq h$  and  $\leq h-1$ , and with  $t$  and  $n-1-t$  elements, respectively,  $0 \leq t \leq n-1$ .

Therefore, summing up possible  $t$ 's, we have

$$f_{\leq h}^n = \sum_{t=0}^{n-1} f_{\leq h}^t f_{\leq h-1}^{n-1-t}.$$

□

## 4.2 Recurrence formula 2

**Theorem 4.2.** *For  $n \geq 2$  and  $h \geq 1$ , we have*

$$f_{\leq h}^n = \sum_{k=1}^{\lfloor \frac{h+2}{2} \rfloor} (-1)^{k-1} \binom{h+2-k}{k} f_{\leq h}^{n-k}. \quad (4.2)$$

*Proof.* Let us prove this theorem by first defining bad elements of a semiorder and then considering removing one or more bad elements from a given semiorder.

**Definition 4.3.** We say an element of a semiorder is *bad*, if the following two conditions hold.

- It is on the first level, or it is smaller than all elements on the level immediately above it, and
- it is on the last level, or it is not larger than any element on the level immediately below it.

*Remark 4.4.* By the above definition, there are no two adjacent levels which both have bad elements. In addition, by Proposition 2.4, if two bad elements are on the same level, they must be equivalent. Therefore there is at most one non-equivalent bad element on each level. We thus only consider one bad element on each level.

**Proposition 4.5.** *For every semiorder, there exist bad elements.*

*Proof.* Assume to the contrary that there is no bad element for some semiorder. Then for an element  $a$  on the last level, there must be some element  $b$  on the last but one level which is not larger than  $a$ . Since the semiorder is  $(\mathbf{2} + \mathbf{2})$ -free,  $b$  must be no larger than any element on the last level. Since  $b$  is not bad, it must not be on the first level, and there must be some element  $c$  on the level immediately above  $b$  that is not larger than  $b$ . Again in order to avoid the  $(\mathbf{2} + \mathbf{2})$ -structure,  $c$  must be no larger than any element on the level immediately below the level  $c$  is on. Continuing, we can find an element on each level that is not larger than any element on the level immediately below it. Since the length of the semiorder is finite, we can finally obtain an element  $d$  on the first level such that  $d$  is not larger than any element on the second level. Then  $d$  is bad, and we get a contradiction. Hence, for every semiorder, bad element exists. □

We are now ready to deduce the recurrence formula (4.2). For fixed  $k$ , consider adjoining  $k$  bad elements to an  $(n-k)$ -element semiorder to get a new semiorder. Specifically, for an  $(n-k)$ -element semiorder of length at most  $h$ , and for  $k$  nonadjacent levels among levels  $1, 2, \dots, h+1$ , we consider adjoining  $k$  elements onto the given levels of the semiorder in the following manner. Say we are adjoining an element onto the  $l^{\text{th}}$  level.

- If  $l = 1$ , let the new element be larger than all elements on the  $i^{th}$  level,  $i \geq 3$ , and be not comparable with any other element.
- If  $l \geq 2$ , and the  $(l - 1)^{th}$  level originally has at least one element, we let the new element be smaller than all elements on the  $(l - 1)^{th}$  level, be larger than all elements on the  $i^{th}$  level,  $i \geq l + 2$ , and be not comparable with any other element.
- If  $l \geq 2$ , and the  $(l - 1)^{th}$  level originally has no element, we hang the new element on the  $l^{th}$  level. We then call the new semiorder an invalid semiorder, and if some semiorder  $r_0$  can be obtained from the invalid semiorder by taking out the hanging elements, we call the invalid semiorder the disguise of  $r_0$ .

By Definition 4.3, in the above adjoining, all new elements which are not hung are bad elements in the new semiorder. There are  $\binom{h+2-k}{k}$  ways to choose  $k$  nonadjacent levels among levels  $1, 2, \dots, h+1$ . Therefore, including multiplicity and the invalid ones, we can obtain  $\binom{h+2-k}{k} \cdot f_{\leq h}^{n-k}$   $n$ -element semiorders from  $(n - k)$ -element semiorders by adjoining  $k$  elements. Let  $\mathcal{S}^k$  be the set of all such  $n$ -element semiorders, including multiplicity. Then  $|\mathcal{S}^k| = \binom{h+2-k}{k} \cdot f_{\leq h}^{n-k}$ .

Let  $\mathcal{R}$  be the set of all  $n$ -element semiorders of length at most  $h$ , and  $\mathcal{R}'$  be the set of all semiorders with at most  $n - 1$  elements and length at most  $h$ . By Proposition 4.5, every semiorder has bad elements, so every semiorder  $r \in \mathcal{R}$  can be obtained by the above process from some  $(n - k)$ -element semiorder of length at most  $h$  and  $k$  given nonadjacent levels, i.e.,  $r \in \mathcal{S}^k$ , for some  $1 \leq k \leq \lfloor \frac{h+2}{2} \rfloor$ . However,  $r$  might be in  $\mathcal{S}^k$  for multiple  $k$ 's, and  $r$  may have multiple copies in  $\mathcal{S}^k$ . Meanwhile,  $\mathcal{S}^k$  may contain some semiorders not in  $\mathcal{R}$ , but are the disguises of some semiorders  $r' \in \mathcal{R}'$ . Notice that  $|\mathcal{R}| = f_{\leq h}^n$ . In the following argument, we calculate the number of copies of a semiorder in each  $\mathcal{S}^k$  and obtain a formula connecting  $|\mathcal{R}|$  and  $|\mathcal{S}^k|$ ,  $1 \leq k \leq \lfloor \frac{h+2}{2} \rfloor$ .

For a semiorder  $r_0 \in \mathcal{R} \cup \mathcal{R}'$ , let  $\mathcal{S}_{r_0}^k$  be the set of all semiorders in  $\mathcal{S}^k$  which are equal to  $r_0$  or a disguise of  $r_0$ . Then  $\mathcal{S}^k = \bigcup_{r \in \mathcal{R} \cup \mathcal{R}'} \mathcal{S}_r^k$ , and

$$\sum_{r \in \mathcal{R} \cup \mathcal{R}'} |\mathcal{S}_r^k| = |\mathcal{S}^k| = \binom{h+2-k}{k} \cdot f_{\leq h}^{n-k}. \quad (4.3)$$

Next, we show that  $\sum_{k=1}^{\lfloor \frac{h+2}{2} \rfloor} (-1)^{k-1} |\mathcal{S}_r^k| = 1$ , for every  $r \in \mathcal{R}$ , and  $\sum_{k=1}^{\lfloor \frac{h+2}{2} \rfloor} (-1)^{k-1} |\mathcal{S}_{r'}^k| = 0$ , for every  $r' \in \mathcal{R}'$ .

1. For a semiorder  $r \in \mathcal{R}$ , assume  $r$  has  $m$  bad elements. Since we adjoined  $k$  elements to an  $(n - k)$ -element semiorder to obtain the semiorder  $r$ , which has  $n$  elements, the  $k$  new elements should all be added to the levels among the  $m$  levels where the bad elements are, and no new element is hung. So  $k \leq m$ . Further notice that for a given  $k$ ,  $1 \leq k \leq m$ , and given  $k$  levels among the  $m$  levels where the bad elements are, there is a unique  $(n - k)$ -element semiorder can be used to adjoin  $k$  bad elements to the chosen levels to obtain semiorder  $r$ . There are  $\binom{m}{k}$  ways to choose the  $k$  levels, so  $|\mathcal{S}_r^k| = \binom{m}{k} \cdot 1$ , and

$$\sum_{k=1}^{\lfloor \frac{h+2}{2} \rfloor} (-1)^{k-1} |\mathcal{S}_r^k| = \sum_{k=1}^m (-1)^{k-1} |\mathcal{S}_r^k| = \sum_{k=1}^m (-1)^{k-1} \binom{m}{k} \cdot 1 = 1. \quad (4.4)$$

2. For a semiorder  $r' \in \mathcal{R}'$ , assume  $r'$  has  $n'$  elements,  $m'$  of which are bad. Further assume that semiorder  $r'$  has length  $h'$ . For a given  $k$ ,  $1 \leq k \leq \lfloor \frac{h+2}{2} \rfloor$ , if we adjoined  $k$  elements to an  $(n - k)$ -element semiorder to obtain  $r'$ , we need to adjoin  $t' = n' - n + k$  elements to levels where the bad elements of  $r'$  are, and hang the remaining  $n - n'$  elements. Moreover, since we hung  $n - n'$

elements, there should be at least  $n - n'$  nonadjacent levels among levels  $h' + 2, h' + 3, \dots, h + 1$ . As a result,  $\left\lceil \frac{h-h'}{2} \right\rceil \geq n - n'$ .

To obtain the semiorder  $r'$ , if we are given  $t'$  levels among the  $m'$  levels where the bad elements are and  $n - n'$  nonadjacent levels among levels  $h' + 2, h' + 3, \dots, h + 1$ , there is a unique  $(n - k)$ -element semiorder to which we can adjoin  $k$  elements to the chosen levels to obtain the disguise of  $r'$ . Notice that there are  $\binom{m'}{t'}$  ways to choose  $t'$  levels among the  $m'$  levels, and  $\binom{h-h'+1-(n-n')}{n-n'}$  ways to choose  $n - n'$  nonadjacent levels from levels  $h' + 2, h' + 3, \dots, h + 1$ . Thus,

$$|\mathcal{S}_{r'}^k| = |\mathcal{S}_{r'}^{t'-n'+n}| = \binom{m'}{t'} \cdot \binom{h-h'+1-(n-n')}{n-n'} \cdot 1.$$

Then

$$\begin{aligned} \sum_{k=1}^{\lfloor \frac{h+2}{2} \rfloor} (-1)^{k-1} |\mathcal{S}_{r'}^k| &= \sum_{t'=0}^{m'} (-1)^{t'-n'+n-1} |\mathcal{S}_{r'}^{t'-n'+n}| \\ &= \sum_{t'=0}^{m'} (-1)^{t'-n'+n-1} \binom{m'}{t'} \cdot \binom{h-h'+1-(n-n')}{n-n'} \cdot 1 \\ &= (-1)^{n-n'-1} \cdot \binom{h-h'+1-(n-n')}{n-n'} \sum_{t'=0}^{m'} (-1)^{t'} \binom{m'}{t'} = 0. \end{aligned} \quad (4.5)$$

However, we should be careful with the special case when  $t' \geq 1$  and the  $(h' + 1)^{th}$  level of  $r'$  has bad elements. When we choose  $t'$  levels among the  $m'$  levels where bad elements are and  $n - n'$  nonadjacent levels among levels  $h' + 2, h' + 3, \dots, h + 1$ , it is possible that the  $(h' + 1)^{th}$  and  $(h' + 2)^{th}$  levels are both chosen. This case should not occur when we directly choose  $k$  nonadjacent levels among levels  $1, 2, \dots, h + 1$ . So we need to take out the overcounts, and thus in this case,

$$\begin{aligned} |\mathcal{S}_{r'}^k| &= |\mathcal{S}_{r'}^{t'-n'+n}| \\ &= \binom{m'}{t'} \cdot \binom{h-h'+1-(n-n')}{n-n'} \cdot 1 - \binom{m'-1}{t'-1} \cdot \binom{h-h'-1-(n-n'-1)}{n-n'-1} \cdot 1 \end{aligned}$$

By similar calculations, we have  $\sum_{k=1}^{\lfloor \frac{h+2}{2} \rfloor} (-1)^{k-1} |\mathcal{S}_{r'}^k| = 0$ .

To conclude the proof, by equations (4.3), (4.4), and (4.5), we have

$$\begin{aligned} f_{\leq h}^n = |\mathcal{R}| &= \sum_{r \in \mathcal{R}} 1 + \sum_{r' \in \mathcal{R}'} 0 = \sum_{r \in \mathcal{R}} \sum_{k=1}^{\lfloor \frac{h+2}{2} \rfloor} (-1)^{k-1} |\mathcal{S}_r^k| + \sum_{r' \in \mathcal{R}'} \sum_{k=1}^{\lfloor \frac{h+2}{2} \rfloor} (-1)^{k-1} |\mathcal{S}_{r'}^k| \\ &= \sum_{k=1}^{\lfloor \frac{h+2}{2} \rfloor} (-1)^{k-1} \sum_{r \in \mathcal{R} \cup \mathcal{R}'} |\mathcal{S}_r^k| \\ &= \sum_{k=1}^{\lfloor \frac{h+2}{2} \rfloor} (-1)^{k-1} \binom{h+2-k}{k} \cdot f_{\leq h}^{n-k}. \end{aligned}$$

□

## 5 The Number of Semiorders of Small Length

We can substitute certain lengths  $H$  in the explicit formulas for the number of semiorders. Though the original formulas are very complicated, we can get some simple results for small values of  $H$ . In this section, we list these simple results and give bijective proofs, which present a clearer view of the number of fixed-length semiorders.

### 5.1 $f_{\leq 1}^n$ , the number of nonisomorphic unlabeled $n$ -element semiorders of length at most one

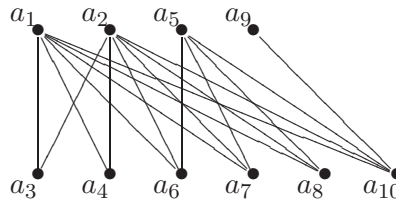
**Theorem 5.1.** For  $n \geq 1$ ,  $f_{\leq 1}^n = 2^{n-1}$ .

We give a simple bijective proof here.

**Proposition 5.2.** For  $n$  elements  $a_1, a_2, \dots, a_n$ , put  $a_1$  on the upper level, and each of  $a_2, \dots, a_n$  either on the lower or upper level. Define the order relations in the following way:  $a_i > a_j$  if and only if  $i < j$ , and  $a_i$  is on the upper level while  $a_j$  is on the lower level.

We claim that the above defines a bijective map from (a) an arrangement of  $n-1$  elements onto two levels in (b) an  $n$ -element semiorder of length at most one.

Here is an example of the map. Say  $n = 10$ , and for  $a_2, \dots, a_{10}$ , let  $\{a_2, a_5, a_9\}$  be on the upper level, and  $\{a_3, a_4, a_6, a_7, a_8, a_{10}\}$  on the lower level. Then the corresponding semiorder looks like:



*Proof.* We first show that the map gives a semiorder of length at most one. It suffices to show that the poset the map gives is indeed a semiorder. Then the only possible violation is a  $(\mathbf{2} + \mathbf{2})$ -structure. If there exist four distinct elements  $a_i, a_j, a_k, a_m$ , such that  $a_i > a_j$ ,  $a_k > a_m$ ,  $a_i \sim a_k$ ,  $a_i \sim a_m$ ,  $a_k \sim a_j$ ,  $a_m \sim a_j$ , then  $a_i, a_k$  must be on the upper level, while  $a_j, a_m$  must be on the lower level, and  $i < j$ ,  $k < m$ . Since  $a_i \sim a_m$ , we must have  $i > m$ ; since  $a_k \sim a_j$ , we must have  $k > j$ . Then  $k > j > i > m > k$ , which is a contradiction. Hence the map gives a semiorder of length at most one.

We then claim that the inverse map is also well-defined, and thus the map is bijective. For a given  $n$ -element  $r$  semiorder with  $m$  elements on the upper level, let  $\rho(r) = (r_1, r_2, \dots, r_n)$ . Then  $r_{m+1} = r_{m+2} = \dots = r_n = 0$ . Say element  $a_{t_i}$  corresponds to  $r_i$ ,  $1 \leq i \leq m$ , and then there should be exactly  $r_i$  elements on the lower level such that their subscripts are larger than  $t_i$ . As a result, note that  $a_1$  is on the upper level, we should also have  $a_{r_1-r_2+2}, a_{r_1-r_3+3}, \dots, a_{r_1-r_m+m}$  on the upper level. In other words, for a given semiorder of length at most one, the elements arranged on the upper level are uniquely determined. Therefore, the inverse map is well-defined.  $\square$

As an example, if we have  $(r_1, r_2, \dots, r_n) = (6, 6, 4, 1, 0, 0, 0, 0, 0, 0)$ , then the elements on the upper level must be  $a_1, a_2, a_5, a_9$ .

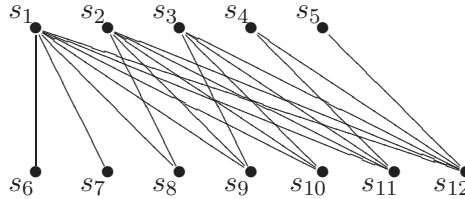
There are  $2^{n-1}$  ways to arrange elements  $a_2, \dots, a_n$  on either upper or lower level, and thus there are  $2^{n-1}$  nonisomorphic unlabeled  $n$ -element semiorders of length at most one.

## 5.2 The number of nonisomorphic trees derived from semiorders of length at most one

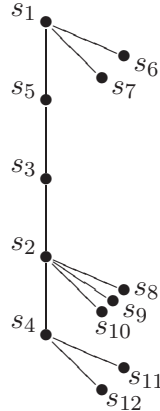
In this subsection, we take a closer look at the unlabeled semiorders of length at most one. For an  $n$ -element semiorder  $S$  of length at most one, and exactly  $m$  elements on the first level, let  $\rho(S) = (r_1, r_2, \dots, r_m, 0, 0, \dots, 0)$ , where there are  $n - m$  0's and  $n - m \geq r_1 \geq r_2 \geq \dots \geq r_m \geq 0$ . Let the elements of the semiorder be  $s_1, s_2, \dots, s_n$ , with  $s_i$  corresponding to  $r_i$ ,  $1 \leq i \leq m$ , and then  $s_1, s_2, \dots, s_m$  are on the upper level.

For a permutation  $\sigma = (a_1, a_2, \dots, a_m)$  of  $\{1, 2, \dots, m\}$ , if we add the relations  $s_{a_1} > s_{a_2} > \dots > s_{a_m}$  to the original semiorder, we get a tree with the main trunk  $s_{a_1} > s_{a_2} > \dots > s_{a_m}$ , and the elements  $s_{m+1}, s_{m+2}, \dots, s_n$  attached to one of the elements on the main trunk in the following manner. For  $m + 1 \leq i \leq n$ , element  $s_i$  is attached to element  $s_j$  on the main trunk if and only if  $s_i < s_j$ , and  $s_i$  is incomparable with all elements on the main trunk that are below  $s_j$ ,  $1 \leq j \leq m$ . We denote the tree derived from semiorder  $S$  and permutation  $\sigma$  by  $T(S, \sigma)$ .

For example, the Hasse diagram of the semiorder  $S$  with  $\rho(S) = (7, 5, 4, 2, 1, 0, 0, 0, 0, 0, 0, 0)$  is as follows:



Suppose that  $\sigma = (1, 5, 3, 2, 4)$ , and then we add the relations  $s_1 > s_5 > s_3 > s_2 > s_4$  to the original semiorder. Then  $T(S, \sigma)$  is



Here we call  $s_1 > s_5 > s_3 > s_2 > s_4$  the main trunk, and say elements  $s_6, s_7$  are attached to  $s_1$ , elements  $s_8, s_9, s_{10}$  are attached to  $s_2$ , and elements  $s_{11}, s_{12}$  are attached to  $s_4$ .

The idea of transforming a semiorder of length at most one to a tree is suggested by R. Stanley, in the context of finding the number of linear extensions of  $n$ -element semiorders of length at most one. Though this idea may not be useful in its original context, we can give a different application.

**Theorem 5.3.** *Given an  $n$ -element semiorder  $R_m$  of length at most one and exactly  $m$  elements on the first level, let  $\rho(R_m) = (r_1, r_2, \dots, r_n)$ . If  $r_i \neq r_j$  for any  $1 \leq i < j \leq m$ , then the number of nonisomorphic unlabeled trees in  $\{T(R_m, \sigma) | \sigma \in S_m\}$  is the Catalan number  $C_m$ .*

*Proof.* A permutation  $\sigma = (a_1, a_2, \dots, a_m)$  of  $\{1, 2, \dots, m\}$  uniquely determines the main trunk. For  $m + 1 \leq i \leq n$ , assume element  $s_i$  is smaller than  $t_i$  elements. Then  $s_i$  must be smaller than  $s_1, s_2, \dots, s_{t_i}$  and is not comparable with the other elements. Let  $s_{u_i}$  be the lowest element on the

main trunk among  $s_1, s_2, \dots, s_{t_i}$ . Then  $s_i$  is attached to  $s_{u_i}$  as a leaf, meaning  $s_i$  is smaller than  $s_{u_i}$  but not comparable with any element on the main trunk below  $s_{u_i}$ .

As a result, for an element  $s_u$  on the main trunk,  $s_u$  has leaves only if it is the lowest element on the main trunk among  $s_1, s_2, \dots, s_t$ , for some  $1 \leq t \leq m$ . In other words, assume  $b_1 < b_2 < \dots < b_k$  are the set of right-to-left minima of the permutation  $\sigma = (a_1, a_2, \dots, a_m)$ , and then only the elements  $s_{b_1}, s_{b_2}, \dots, s_{b_k}$  may have leaves. Further notice that the numbers of leaves attached to  $s_{b_1}, s_{b_2}, \dots, s_{b_k}$  are  $r_{b_1} - r_{b_2}, r_{b_2} - r_{b_3}, \dots, r_{b_{k-1}} - r_{b_k}, r_{b_k}$ , respectively, and  $r_i \neq r_j$  for any  $1 \leq i < j \leq m$ . Therefore, for a given  $n$ -element semiorder  $R_m$  of length at most one and exactly  $m$  elements on the first level, the value and position of the right-to-left minima of the permutation  $\sigma = (a_1, a_2, \dots, a_m)$  uniquely determines  $T(R_m, \sigma)$ .

For instance, in the example above, we have  $\rho(R_5) = (r_1, r_2, \dots, r_{12}) = (7, 5, 4, 2, 1, 0, 0, 0, 0, 0, 0, 0)$  and  $\sigma = (1, 5, 3, 2, 4)$ . The right-to-left-minima of  $\sigma$  and their positions with  $\sigma$  is given by  $1, *, *, 2, 4$ . Then the corresponding tree  $T(R_5, \sigma)$  has five nodes on the main trunk,  $r_1 - r_2 = 2$  leaves attached to the first node,  $r_2 - r_4 = 3$  leaves attached to the forth node, and  $r_4 = 2$  leaves attached to the fifth node.

Therefore, the number of all possible nonisomorphic unlabeled trees in  $\{T(R_m, \sigma) | \sigma \in S_m\}$  is the number of ways to specify the values and positions of the right-to-left minima of permutations  $\sigma \in S_m$ . That is, if we let  $RtLM(\sigma) = \{(a, \sigma(a)) | 1 \leq a \leq m, \sigma(a) \text{ is a right-to-left minima in } \sigma\}$ , then  $\#\{T(R_m, \sigma) | \sigma \in S_m\} = \#\{RtLM(\sigma) | \sigma \in S_m\}$ . We calculate  $\#\{RtLM(\sigma) | \sigma \in S_m\}$  in the following lemma.

**Lemma 5.4.** *Let  $RL(\sigma)$  be the number of right-to-left minima of the permutation  $\sigma$ . For  $1 \leq k \leq m$ , let  $f(m, k) = \#\{RtLM(\sigma) | \sigma \in S_m, RL(\sigma) = k\}$ . Then  $f(m, k) = N(m, k) = \frac{1}{m} \binom{m}{k} \binom{m}{k-1}$ , a Narayana number.*

For example, for  $m = 3$  and  $k = 2$ ,  $f(3, 2) = \#\{RtLM(\sigma) | \sigma \in S_3, RL(\sigma) = 2\} = \#\{RtLM(\sigma) | \sigma = (23), (12), \text{ or } (132)\} = \#\{\{(3, 2), (1, 1)\}, \{(3, 3), (2, 1)\}, \{(3, 2), (2, 1)\}\} = 3$ .

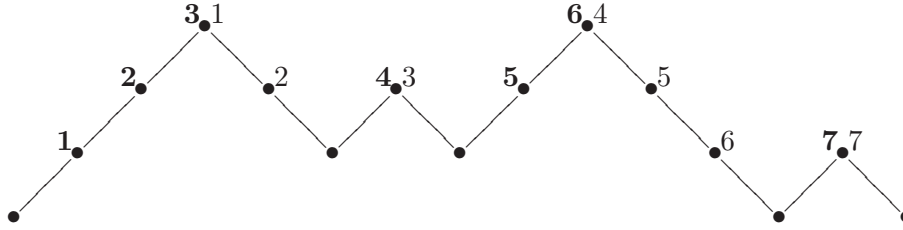
*Proof.* For  $1 \leq k \leq m$ , the Narayana number  $N(m, k)$  is equal to the number of Dyck paths of semilength  $m$  with  $k$  peaks, which are the turning points from a  $(1, 1)$  step to a  $(1, -1)$  step on the path. We prove the lemma by establishing a bijection between (i) Dyck paths of semilength  $m$  with  $k$  peaks, and (ii) the collection of different  $RtLM(\sigma)$ 's for  $\sigma \in S_m, RL(\sigma) = k$ . We define a map from (i) to (ii) as follows:

Given a Dyck path of semilength  $m$  with  $k$  peaks, let us read the Dyck path from left to right and do the following:

- Label the endpoints of  $(1, 1)$  steps from left to right with 1 to  $m$ . Since there are  $m$   $(1, 1)$  steps, there should be  $m$  such endpoints.
- Label the startpoints of  $(1, -1)$  steps from left to right with 1 to  $m$ . Since there are  $m$   $(1, -1)$  steps, there should be  $m$  such startpoints.
- Notice that a point on the Dyck path is a peak if and only if it is both an endpoint of a  $(1, 1)$  step and a startpoint of a  $(1, -1)$  step. Let  $(i, j)$  be the coordinate of a peak, if the peak is the  $i^{th}$  endpoint and the  $j^{th}$  startpoint. Assume the coordinate of the  $k$  peaks are  $(a_1, b_1), (a_2, b_2), \dots, (a_k, b_k)$ . Then  $a_1 < a_2 < \dots < a_k = m$  and  $1 = b_1 < b_2 < \dots < b_k$ .
- Obtain a specification of the values and positions of the right-to-left minima of a permutation by putting the number  $b_i$  on position  $a_i$ ,  $1 \leq i \leq k$ .

**Proposition 5.5.** *The above map is valid, i.e.,  $\{(a_i, b_i), 1 \leq i \leq k\} = \text{RtLM}(\sigma)$ , for some  $\sigma \in S_m$ ,  $\text{RL}(\sigma) = k$ , and the above map is a bijection.*

For example, if we have the following Dyck path:



The semilength of the Dyck path is  $m = 7$ , and it has 4 peaks. The numbers in bold face are the labels for the endpoints of  $(1, 1)$  steps, and the numbers in ordinary type are the labels for the startpoints of  $(-1, 1)$  steps. Then the coordinates of the 4 peaks are  $(3, 1), (4, 3), (6, 4), (7, 7)$ . We put 1 on position 3, 3 on position 4, 4 on position 6, 7 on position 7, and then we get a possible specification of the values and positions of the right-to-left minima of a permutation:

$$*, *, 1, 3, *, 4, 7. \quad (5.1)$$

*Proof.* • We first show that the above map gives us a valid specification of the values and positions of the right-to-left minima of some permutation  $\sigma \in S_m$ . We prove the validity by constructing such  $\sigma$ .

Assume the labels of the  $m - k$  endpoints of the  $(1, 1)$  steps which are not peaks, are  $c_1 < c_2 < \dots < c_{m-k}$ ; assume the labels of the  $m - k$  startpoints of the  $(1, -1)$  steps which are not peaks are  $d_1 < d_2 < \dots < d_{m-k}$ . Then  $\{a_1, a_2, \dots, a_k, c_1, c_2, \dots, c_{m-k}\} = \{b_1, b_2, \dots, b_k, d_1, d_2, \dots, d_{m-k}\} = \{1, 2, \dots, m\}$ . Since the path never goes below the  $x$ -axis, we must have  $c_j < d_j$  for every  $1 \leq j \leq m - k$ .

Let  $\sigma$  be a permutation such that  $\sigma(a_i) = b_i$ ,  $1 \leq i \leq k$ , and  $\sigma(c_j) = d_j$ ,  $1 \leq j \leq m - k$ . We will show that  $\text{RtLM}(\sigma) = \{(a_i, b_i), 1 \leq i \leq k\}$ .

For every  $j$ ,  $1 \leq j \leq m - k$ , since the Dyck path never goes below the  $x$  axis, there must be a peak  $i$  on the path between  $c_j$  and  $d_j$ . Then  $c_j < a_i$  and  $b_i < d_j$ , and thus in  $\sigma$ ,  $b_i = \sigma(a_i)$  is smaller than  $d_j = \sigma(c_j)$ , while  $a_i > c_j$ , i.e.,  $b_i$  is to the right of  $d_j$ . Therefore  $d_j$  cannot be a right-to-left minimum.

On the other hand, for every  $i$ ,  $1 \leq i \leq k$ , there does not exist some  $1 \leq i' \leq k$  such that  $b_{i'} < b_i$  and  $a_{i'} > a_i$ . In addition, if there exists some  $1 \leq j \leq m - k$  such that  $d_j = \sigma(c_j) < b_i = \sigma(a_i)$  and  $c_j < a_i$ , then by the above paragraph there exists  $1 \leq i' \leq k$ , such that  $b_{i'} < d_j < b_i$  and  $a_{i'} > c_j > a_i$ . We obtain a contradiction. As a result,  $\text{RtLM}(\sigma) = \{(a_i, b_i), 1 \leq i \leq k\}$ .

For example, let us construct the permutation  $\sigma$  for the above example:  $c_1 = 1, c_2 = 2, c_3 = 5$  and  $d_1 = 2, d_2 = 5, d_3 = 6$ . We obtain  $\sigma = (2, 5, 1, 3, 6, 4, 7)$ , and this permutation exactly corresponds to the right-to-left minima as shown in Example 5.1.

• We now show that the inverse of the map is well-defined, and thus the map is bijective.

For a given specification  $\{(a_i, b_i), 1 \leq i \leq k\} = \text{RtLM}(\sigma)$ , for some  $\sigma \in S_m$ , we have  $a_1 < a_2 < \dots < a_k = m$  and  $1 = b_1 < b_2 < \dots < b_k$ . We construct the corresponding Dyck path as follows: when we walk along the path from left to right, we first walk up  $a_1$  steps



and then turn down, and walk down  $b_2 - b_1$  steps and then turn up. We continue to walk up  $a_2 - a_1$  steps and then turn down, and walk down  $b_3 - b_2$  steps and then turn up. In general, we walk up  $a_i - a_{i-1}$  steps and then turn down, and walk down  $b_{i+1} - b_i$  steps,  $2 \leq i \leq m-1$ . In the end, we walk up  $a_m - a_{m-1}$  steps and walk down  $m+1 - b_m$  steps. During the walk, we walk up in total  $a_1 + (a_2 - a_1) + \cdots + a_m - a_{m-1} = a_m = m$  steps, and walk down in total  $(b_2 - b_1) + \cdots + (b_m - b_{m-1}) + m+1 - b_m = m+1 - b_1 = m$  steps, and we make  $k$  turns from up to down. Since  $\{(a_i, b_i), 1 \leq i \leq k\}$  is a collection of right-to-left minima of some permutation, for any  $1 \leq i \leq k-1$ ,  $a_i \leq a_{i+1} - 1 \leq b_{i+1} - 1 = b_{i+1} - b_1$ , so we never walk below the  $x$ -axis on the path. Therefore we get a unique Dyck path with semilength  $m$  and  $k$  peaks. Hence the inverse map is well-defined, so Lemma 5.4 is proved.  $\square$

 $\square$  $\square$ 

To conclude the proof of Theorem 5.3, it is well-known that the Catalan number  $C_m = \sum_{k=1}^m N(m, k)$ , so we have

$$\begin{aligned} & \text{The number of nonisomorphic unlabeled trees in } \{T(R_m, \sigma) | \sigma \in S_m\} \\ &= \#\{RtLM(\sigma) | \sigma \in S_m\} \\ &= \sum_{k=1}^m f(m, k) = \sum_{k=1}^m N(m, k) = C_m = \frac{1}{m+1} \binom{2m}{m}. \end{aligned}$$

 $\square$ 

*Remark 5.6.* Theorem 5.3, along with Lemma 5.4, gives another combinatorial explanation of the Catalan number.

### 5.3 The generating function for the number of non-isomorphic labeled $n$ -element semiorders of length at most one

Recall that an ordered partition of a set is a partition of the set into some pairwise disjoint nonempty subsets, together with a linear ordering of these subsets. From the generating function (3.4) for  $G_{\leq h}(x)$ , we get  $G_{\leq 1}(x) = \frac{1-e^{-x}}{2e^{-x}-1} = \frac{e^x-1}{2-e^x}$ , which is exactly the exponential generating function for the number of ordered partitions [5, p.472]. As a result, we can get the following theorem:

**Theorem 5.7.** *The number of non-isomorphic labeled  $n$ -element semiorders of length at most one is equal to the number of ordered partitions of  $[n]$ .*

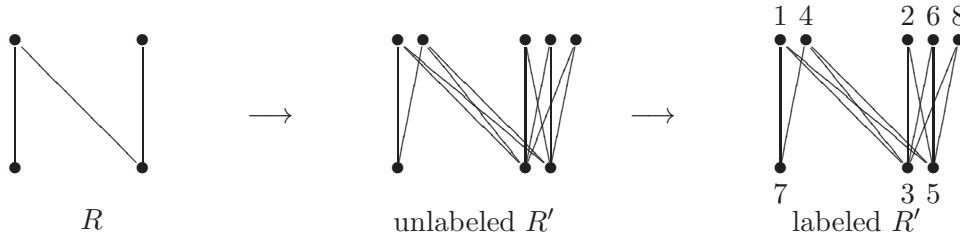
We give a simple bijective proof to Theorem 5.7.

**Proposition 5.8.** *For an ordered partition  $(A_1, \dots, A_k)$  of  $[n]$ , let  $|A_i| = a_i$ ,  $1 \leq i \leq k$ , so  $n = a_1 + a_2 + \cdots + a_k$ . Define the semiorder  $R$  by  $\rho(R) = (m, m-1, \dots, 1, 0, \dots, 0)$ , where  $m = \lfloor \frac{k}{2} \rfloor$ , and there are  $\lceil \frac{k}{2} \rceil$  0's. Then  $R$  has  $k$  elements. Say the elements are  $t_1, t_2, \dots, t_k$ , with  $t_i$  corresponding to the  $i^{\text{th}}$  entry of  $R$ 's integer vector. Let  $R'$  be another semiorder such that its contraction  $c(R')$  (defined in Lemma 3.4) is  $R$ , and in  $R'$ , the sizes of the equivalence classes are  $a_1, a_2, \dots, a_k$ , respectively, with  $a_i$  corresponding to equivalence class  $t_i$ . Label the elements in the  $i^{\text{th}}$  equivalence class with the corresponding numbers in  $A_i$ ,  $1 \leq i \leq k$ .*

*We claim that the above defines a bijective map from ordered partitions of  $[n]$  to  $n$ -element labeled semiorders  $R'$  of length at most one.*



For example, if we have an ordered partition  $\{1, 4\}\{2, 6, 8\}\{7\}\{3, 5\}$ , then  $k = 4$ ,  $m = 2$ , and  $(a_1, a_2, a_3, a_4) = (2, 3, 1, 2)$ . We have  $\rho(R) = (2, 1, 0, 0)$ , and then the map works as follows:



*Proof.* We first show that the map takes an ordered partition of  $[n]$  to a labeled  $n$ -element semiorder of length at most one. The size of every equivalence class of  $R$  is one, and thus semiorder  $R$  is a valid contraction. The length of  $R$  is at most one, and thus  $R'$  also has length at most one. In addition,  $R$  has  $k$  elements, so  $R'$  has  $k$  equivalent classes. Therefore, we can construct the equivalence classes of  $R$  to have sizes  $a_1, a_2, \dots, a_k$ . Moreover, within an equivalence class with  $a_i$  elements,  $1 \leq i \leq k$ , since we only consider nonisomorphic semiorders, it does not matter which of the  $a_i$  numbers in  $A_i$  is assigned to which element in this equivalence class. Hence the way to label elements is unique up to isomorphism. Thus each ordered partition of  $[n]$  uniquely corresponds to a labeled  $n$ -element semiorder of length at most one.

Next we show that the inverse map is well-defined and uniquely determines an ordered partition of  $[n]$ . Let the labeled  $n$ -element semiorder  $R'$  of length at most one have  $k$  equivalence classes. Let us group up the labels within every equivalence class, so

$$[n] = \bigcup \{\text{labels in each equivalence class}\}$$

To obtain an ordered partition of  $[n]$ , it suffices to find the way to order the  $k$  equivalence classes of  $R'$ . The contraction  $c(R')$  must have length at most one with  $\rho(c(R')) = (\lfloor \frac{k}{2} \rfloor, \lfloor \frac{k}{2} \rfloor - 1, \dots, 1, 0, \dots, 0)$ , where there are  $\lceil \frac{k}{2} \rceil$  0's.

Order the  $k$  elements of  $c(R')$  such that the  $i^{\text{th}}$  element corresponds to the  $i^{\text{th}}$  entry of  $\rho(c(R'))$ . Afterwards, we can order the  $k$  equivalence classes of  $R'$  correspondingly. Thus we get a unique ordered partition of  $[n]$ , so the inverse map is well-defined.  $\square$

#### 5.4 The number of non-isomorphic unlabeled $n$ -element semiorders of length at most three

**Theorem 5.9.** For  $n \geq 1$ , we have

$$f_{\leq 3}^n = \frac{3^{n-1} + 1}{2}. \quad (5.2)$$

**Corollary 5.10.** For  $n \geq 2$ ,  $f_3^n = 3f_3^{n-1} + f_{\leq 2}^{n-2} - 1 = 3f_3^{n-1} + f_2^{n-2} + f_1^{n-2}$ .

*Proof.* By Theorem 4.2,  $f_{\leq 2}^n = 3f_{\leq 2}^{n-1} - f_{\leq 2}^{n-2}$ . By equation (5.2),  $f_{\leq 3}^n = 3f_{\leq 3}^{n-1} - 1$ . Therefore,

$$\begin{aligned}
 f_3^n &= f_{\leq 3}^n - f_{\leq 2}^n = 3f_{\leq 3}^{n-1} - 1 - (3f_{\leq 2}^{n-1} - f_{\leq 2}^{n-2}) \\
 &= 3(f_{\leq 3}^{n-1} - f_{\leq 2}^{n-1}) + f_{\leq 2}^{n-2} - 1 \\
 &= 3f_3^{n-1} + f_2^{n-2} + f_1^{n-2}.
 \end{aligned}$$

$\square$

*Remark 5.11.* Theorem 5.9 can be directly derived from equation (3.3), or from the recurrence formula in Theorem 4.2. However, we believe that there should be a more straightforward bijective proof going on, which leaves an open question for this paper.

## References

- [1] N. G. de Bruijn, D. E. Knuth, and S. O. Rice *The average height of planted plane trees*, Graph Theory and Computing, Academic Press, New York, 15–22 (1972).
- [2] J. L. Chandon, J. Lemaire, and J. Pouget, *Dénombrement des quasi-ordres sur un ensemble fini*, Math. Inform. Sci. Humaines 62, 61–80, 83 (1978).
- [3] P. C. Fishburn and W. T. Trotter *Linear extensions of semiorders: A maximization problem*, Discrete Mathematics 103, 25–40 (1992).
- [4] S. Kitaev and J. Remmel *Enumerating  $(2+2)$ -free posets by the number of minimal elements and other statistics*, <http://linux.bucknell.edu/~pm040/Slides/Kitaev.pdf>, 2011.
- [5] R. P. Stanley *Enumerative Combinatorics*, volume 1, second edition, Cambridge University Press, Cambridge, 2011.
- [6] R. L. Wine and J. E. Freund *On the Enumeration of Decision Patterns Involving  $n$  Means*, Ann. Math. Stat. 28, 256–259 (1957).
- [7] Y. X. Zhang *On Poset Enumeration and Automorphisms*, M.I.T. Ph.D. thesis, in preparation.